



A cooperative algorithm for multi-objective optimization: multiple-gradient descent algorithm (MGDA)

Jean-Antoine Désidéri, Andrea Minelli, Adrien Zerbinati

► To cite this version:

Jean-Antoine Désidéri, Andrea Minelli, Adrien Zerbinati. A cooperative algorithm for multi-objective optimization: multiple-gradient descent algorithm (MGDA). 4th Inverse Problems, Design and Optimization Symposium (IPDO-2013), Ecole des Mines d'Albi-Carmaux, Jun 2013, Albi, France. hal-00934637

HAL Id: hal-00934637

<https://hal.inria.fr/hal-00934637>

Submitted on 22 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A COOPERATIVE ALGORITHM FOR MULTI-OBJECTIVE OPTIMIZATION: MULTIPLE-GRADIENT DESCENT ALGORITHM (MGDA)

Jean-Antoine Désidéri^{a,b}, Andrea Minelli^{b,d} and Adrien Zerbinati^{c,d}

^a INRIA, Centre de Sophia Antipolis Méditerranée, 06902 Sophia Antipolis Cedex, France,

^b ONERA - The French Aerospace Lab, F-92190, Meudon, France,

^c BB0143 CSTJF Total, 174 avenue Larribau, 64018 Pau cedex, France,

^d University of Nice - Sophia Antipolis, France,

Jean-Antoine.Desideri@inria.fr Andrea.Minelli@onera.fr Adrien.Zerbinati@total.com

ABSTRACT

The Multiple-Gradient Descent Algorithm (MGDA) has been proposed and tested for the treatment of multi-objective differentiable optimization. Originally introduced in [1], the method has been tested and reformulated in [4]. Its efficacy to identify the Pareto front has been demonstrated in [5], in comparison with an evolutionary strategy. Recently, a variant, MGDA-II, has been proposed in which the descent direction is calculated by a direct procedure [3] based on a Gram-Schmidt orthogonalization process (GSP) with special normalization. This algorithm was tested in the context of a simulation by domain partitioning, as a technique to match the different interface components concurrently [2]. The experimentation revealed the importance of scaling, and a slightly modified normalization procedure was proposed (MGDA-IIb). Two variants have since been proposed. The first, MGDA-III, realizes two enhancements. Firstly, the GSP is conducted incompletely whenever a test reveals that the current estimate of the direction of search is adequate also w.r.t. the gradients not yet taken into account; this improvement simplifies the identification of the search direction when the gradients point roughly in the same direction, and makes the Fréchet derivative common to several objective-functions larger. Secondly, the order in which the different gradients are considered in the GSP is defined in a unique way devised to favor an incomplete GSP. In the second variant, MGDA-IV, the question of scaling is addressed when the Hessians are known. A variant is also proposed in which the Hessians are estimated by the Broyden-Fletcher- Goldfarb-Shanno (BFGS) formula. The method has been successfully applied to a classical test-case proposed by Fonseca [5]. Other examples of application of this method to optimum-shape design in aerodynamics were presented [6].

In this new contribution, the basic principle of the method is recalled. A meta-model-assisted extension is proposed and applied to the shape optimization of a generic supersonic aircraft configuration w.r.t. drag and sonic-boom reduction. This cooperative algorithm permits to identify points on the Pareto set associated with these two objective functions. From one such point, a competitive Nash game with adapted territory splitting can be initiated to identify a path in function space tangent to the Pareto front. Thus, the two approaches, cooperative and competitive algorithms, can be combined to generate quickly a set of designs in the vicinity of the Pareto front.

1 INTRODUCTION

In multi-objective optimization, one classically refers to the notion of Pareto-optimality to evaluate design-points in efficiency. In short, if $Y \in \mathbb{R}^N$ denotes the design-vector, a design-point Y^1 is said to dominate in efficiency the design-point Y^2 iff

$$J_i(Y^1) \leq J_i(Y^2) \quad (1)$$

for all objective-functions $\{J_i\}$ ($i = 1, \dots, n$), and at least one of these inequalities holds strictly. The Pareto set is the set of all design-points dominated by no other, and the Pareto front its image in the function space. This article intends to be a contribution to the identification of the Pareto set by differentiable optimization tools.

2 MULTIPLE GRADIENT DESCENT ALGORITHM (MGDA)

2.1 Basic definition

We consider the problem of simultaneous minimization of n objective functions of N design variables, $J_i(\mathbf{x})$ ($i = 1, \dots, n$; $\mathbf{x} \in \mathbb{R}^N$, design vector). The dimensions n and N are arbitrary, although in many applications $n \leq N$. Let

\mathbf{x}^0 be a particular design-point about which the objective functions are smooth (say C^2 in practice) and locally convex. Denote $u_i^0 = \nabla J_i(\mathbf{x}^0)$ ($i = 1, \dots, n$) the gradients, and define the following convex hull:

$$\bar{U} = \left\{ w \in \mathbb{R}^N / w = \sum_{i=1}^n \alpha_i u_i; \alpha_i \geq 0, \forall i; \sum_{i=1}^n \alpha_i = 1 \right\}. \quad (2)$$

\bar{U} is a closed, bounded and convex set associated in the affine space \mathbb{R}^N with a polyhedron of at most n vertices. Hence \bar{U} admits a unique element of minimum norm, say ω [1]. Two cases are possible:

1. $\omega = 0$, and we say that \mathbf{x}^0 is a point of Pareto-stationarity, a necessary condition for Pareto-optimality;
2. or $\omega \neq 0$, and the directional derivatives of the objective functions satisfy the inequalities:

$$(u_i^0, \omega) \geq \|\omega\|^2; \quad (3)$$

hence, $-\omega$ is a descent direction common to all the objective functions.

In the latter case, we define MGDA as the iteration that uses $-\omega$ as the direction of search, and a step-size adjusted to maximize the smallest absolute decrease of the criteria. Accumulation points of this method are Pareto-optimal designs [1]. In this way, MGDA generalizes to the multi-objective optimization the classical steepest-descent method [10].

In the particular case of two criteria, the minimum-norm vector is known analytically. Figure 1 then shows vector ω in the three different possible cases.

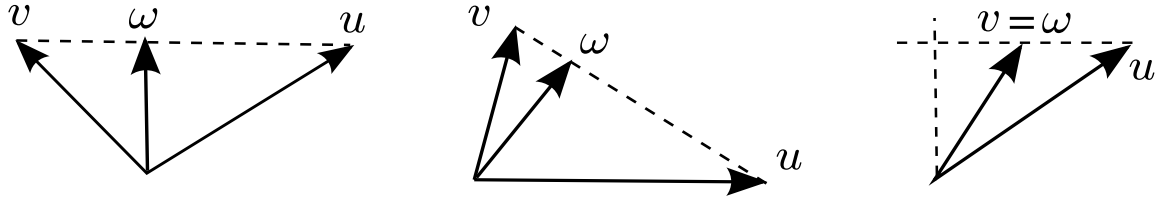


Figure 1: Various possible configurations of the two gradients-vectors $u = u_1$ and $v = u_2$ and the minimal-norm element ω .

As a first illustration of the method, several analytical multi-objective optimization test-cases proposed in [7] have been solved by MGDA [5], and some of these results are presented next in comparison with an evolutionary strategy.

2.2 Analytical validation

The test-case corresponds to the two-objective unconstrained minimization of the following functions :

$$\begin{cases} f_1(\mathbf{x}) = 1 - \exp\left(-\sum_{i=1}^3 \left(x_i \frac{1}{\sqrt{3}}\right)^2\right) \\ f_2(\mathbf{x}) = 1 - \exp\left(-\sum_{i=1}^3 \left(x_i + \frac{1}{\sqrt{3}}\right)^2\right) \end{cases}, \quad \mathbf{x} = (x_1, x_2, x_3). \quad (4)$$

The design variable is $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$. This test-case is known to yield a continuous but non convex Pareto set in function space. The Pareto front was identified by Deb using the well-known genetic algorithm NSGA-II [7].

From a given starting point, MGDA converges quickly (6 steps in this example) and provides an accurately-defined design-point on the Pareto set. After applying the method from a set of 60 initial design-points distributed over a sphere in the design space (see Figure 3), we have obtained an accurate discretization of the known-analytically Pareto set.

2.3 Meta-Model-Assisted Multiple-Gradient Descent Algorithm, MA-MGDA

In PDE-constrained optimization, and in particular in optimum-shape design in aerodynamics, the calculation of function values and their gradients can be very computationally demanding, and usually requires substantial methodological developments. To alleviate this task, in this article, we investigate the possibility of calculating approximate gradients from a surrogate model, or meta-model, devised from a database of high-fidelity function values. In the applications considered presently, the high-fidelity models are associated with 3D compressible flows governed by the Euler or RANS equations.

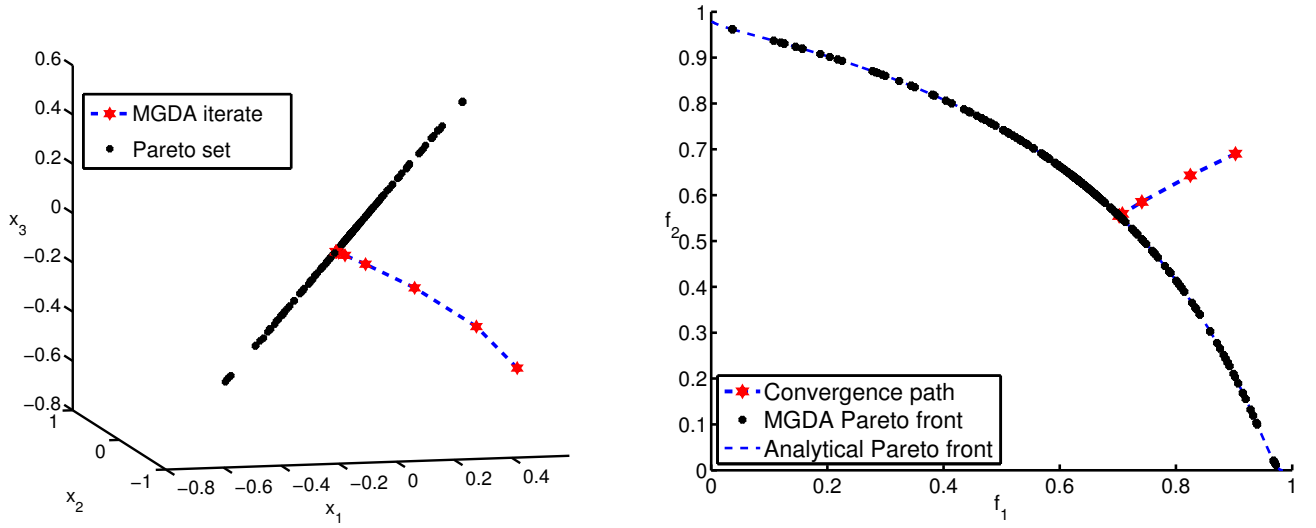


Figure 2: Convergence of MGDA from an initial design point to the non dominated set.

We proceed as follows (see flowchart in Figure 4). An initial set of design points is generated using of a latin hypercube sampling in \mathbb{R}^N . The sampling serves two purposes. Firstly, the function values corresponding to the sampling form a database supporting Kriging meta-models, surrogate of the actual objective functions. Secondly, some of these sampling points are used to initiate independent MGDA iterations applied to the multi-objective minimization of the meta-models, and converging to Pareto-stationary points (associated with the meta-models). These Pareto-stationary points are then evaluated according to the high-fidelity models to enrich the database and proceed with the next update. A filtering method is used additionally to remove points found too close to an existing design-point, in order to avoid redundancy.

3 AERO-ACOUSTIC SHAPE OPTIMIZATION OF A WING-BODY SUPERSONIC CONFIGURATION

As a follow-up of the competitive treatment of two-objective optimization problem in [9], the optimum-shape design of a low-boom/low-drag supersonic business jet design problem is considered here as an application test-case for MA-MGDA. Indeed, sonic boom is one of the main limiting factors to the development of civil supersonic transportation. As the driving design for low-boom is not compliant with the low-drag one, our goal is to provide a trade-off between aerodynamics and acoustics. The MGDA algorithm is adopted to optimize the shape of a SSBJ wing-body configuration at the design condition $M = 1.6$, $AoA = 2$ deg and flight altitude $h_Z = 18,000$ m. The aero-acoustic multi-objective problem can be stated as follows:

$$\text{Minimize } \begin{cases} J_A(\mathbf{x}) = c_D \\ J_B(\mathbf{x}) = \sum \Delta p \end{cases} \quad \text{subject to : } \begin{cases} g = c_L - c_{L_0} \geq 0 & \text{with } c_{L_0} = 0.1 \\ \text{no constraints} \end{cases} \quad (5)$$

where $\mathbf{x} \in \mathbb{R}^{10}$ represents the vector of design variables. The geometrical variables under consideration are lengths, angles and relative distances associated with the configuration depicted in Figure 5. These are denoted $\{h_i\}$, and the actual design variables in the optimization, $\{x_i\}$ (components of \mathbf{x}) are related to them through

$$h_i = \bar{h}_i(1 + x_i\delta_i) \quad \forall i > 1 \quad (6)$$

where \bar{h}_i is the value for the reference configuration, δ_i is a maximum allowable variation. An exception is made for the nose deflection variable since $\bar{h}_1 = 0$; instead $h_1 = 0.5x_1$. The design variables set \mathbf{x} are dimensionless and can assume values between -1 and 1. Table 1 shows the geometrical variables and their respective modification allowed, while parameterization is shown in Figure 5.

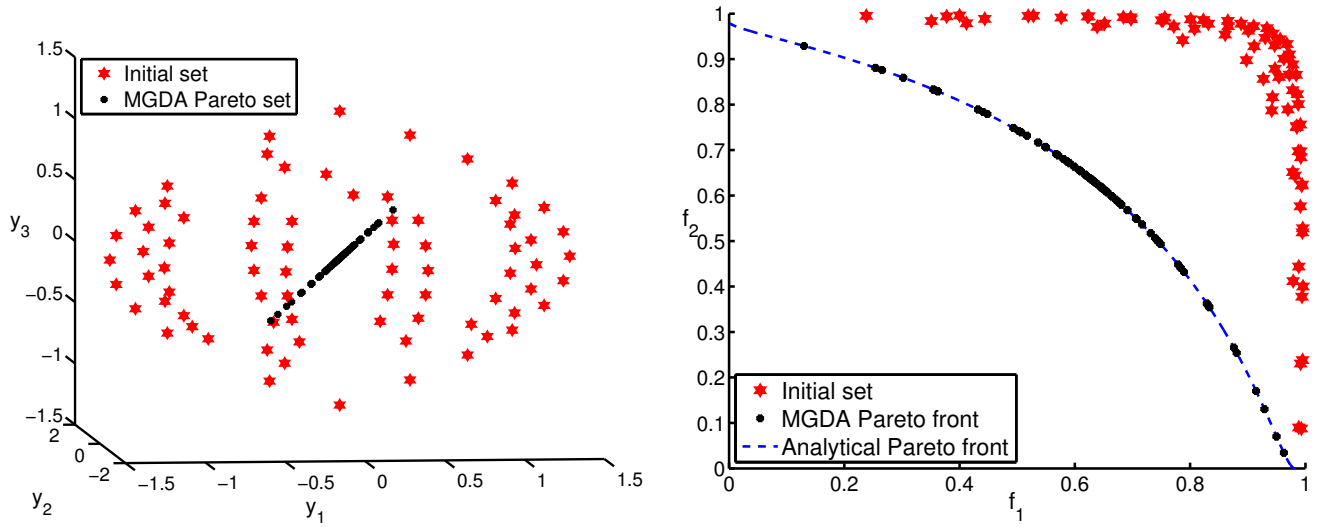


Figure 3: Convergence of MGDA from initial design points around Pareto front, for a classical test case proposed by Fonseca, in design space (left), in function space (right).

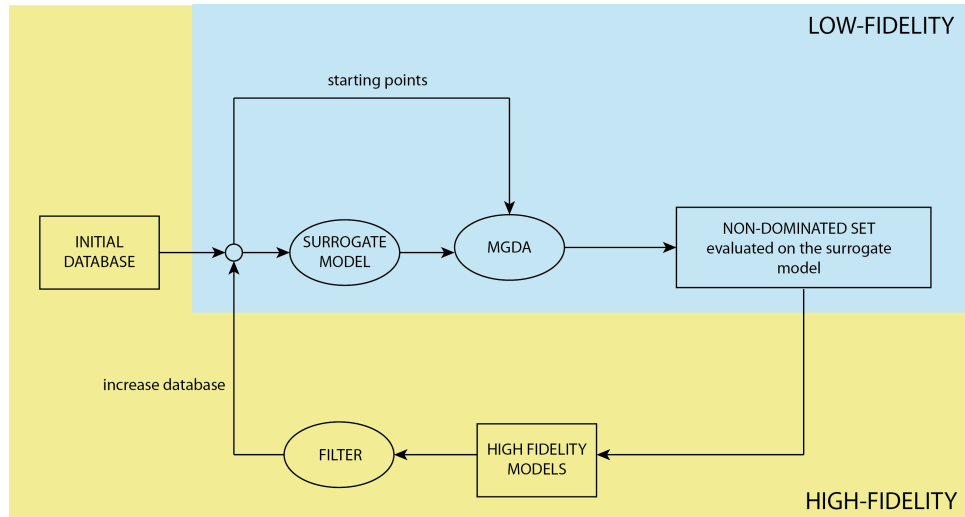


Figure 4: MGDA with surrogate model scheme. A surrogate model based on an initial database is trained. Then, MGDA [1] is applied from each database point. Thus a non dominated set on the surrogate model is obtained.

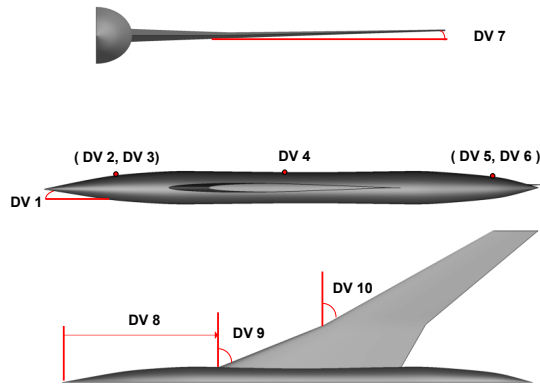


Figure 5: Wing-body parameterization.

Variable	Specification	\bar{h}_i	δ_i
DV1	nose deflection [deg]	0	-
DV2	x-coordinate nose section [m]	4	2%
DV3	radius nose section [m]	0.8	10%
DV4	radius cabin section [m]	1.015	10%
DV5	x-coordinate rear section [m]	28	2%
DV6	radius rear section [m]	0.535	10%
DV7	dihedral angle [deg]	3	100%
DV8	relative wing position [adim]	0.367	10%
DV9	inner wing swept angle [deg]	65	10%
DV10	outer wing swept angle [deg]	56	10%

Table 1: Design variables set definition.

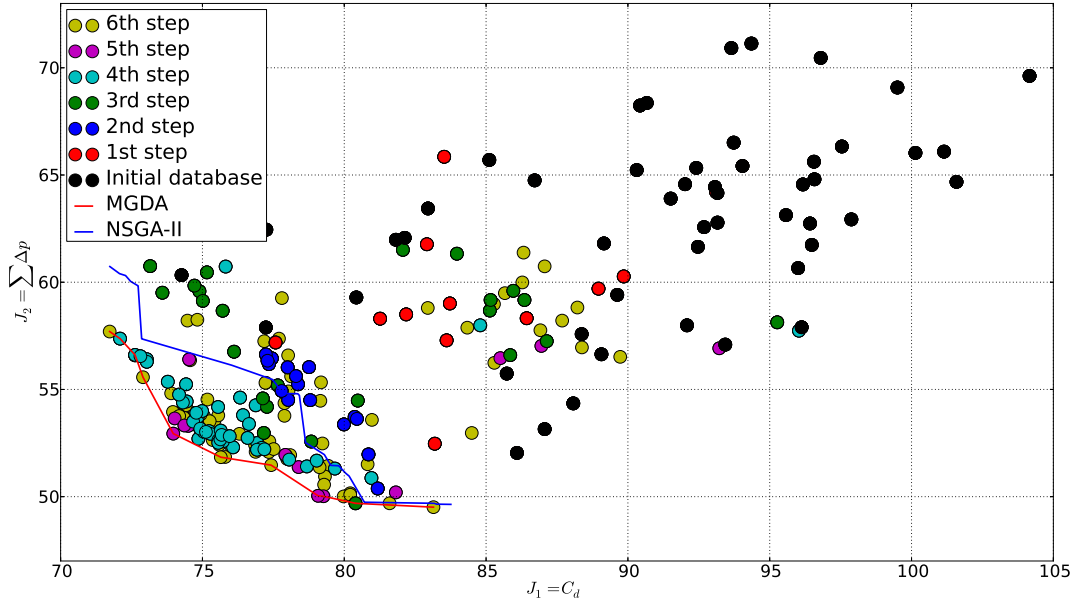


Figure 6: Convergence history at fixed maximum high fidelity function evaluation calls.

The initial Kriging surrogate model has been defined using an LHS database of 100 points. Six MA-MGDA complete iterations are performed, each one including the MGDA convergence on the Kriging model, the evaluation of the Pareto set of solutions with high-fidelity models and the Kriging model update. The MA-MGDA algorithm is able to identify the Pareto front of the high fidelity model with only 166 function evaluations. The NSGA-II algorithm applied on the high fidelity model is used as comparison in order to evaluate the quality of the solution evaluated using MA-MGDA. The comparison is performed at the same number of high-fidelity function evaluations by the high-fidelity model, thus ensuring an identical computational cost. Results clearly show that the design-points in the Pareto front evaluated using MA-MGDA all dominate the solutions obtained by NSGA-II. This demonstrates the efficiency of MA-MGDA, as it is able to converge more accurately or rapidly to the actual Pareto front. Thus the present method is well suited for computationally expensive problems. In this example, the diversity of the solutions is comparable between the two algorithms.

Three different configurations (see table 2) that belong to the Pareto front (figure 6) have been retained and compared in figure 7 and 8.

Configuration	J_1	J_2
A	71.72 dc	57.7 Pa
B	73.98 dc	52.94 Pa
C	83.14 dc	49.51 Pa

Table 2: Selected configurations that belong to the Pareto front.

The low-drag configuration (A) and the trade-off configuration (B) shows minor modifications on the fuselage geometry. The modifications impact mainly the wing plan-form. Configuration (B) shows a reduction of the dihedral angle, with respect to the configuration A that is beneficial for the rear shock of the ground signature (see Fig. 7(b)) with limited deterioration on the aerodynamic performance. In addition the dihedral angle act modifying the acoustic footprint at ground, in particular the duration of the signature. The low boom configuration (C) shows a strong downward displacement of the nose angle, which has as a consequence a modification of the flow condition that reaches the wing leading edge. The expansion before the wing shock is reduced and determines a reduction of the following shock. The front shock is almost not modified in any of the selected geometries. The algorithm has identified the wing and the rear shock reductions as the most promising regions to improve the functions of interest with respect to the chosen parameterization.

Examining closely the pressure distribution just below the aircraft in figure 7(a), it is possible to note that the design

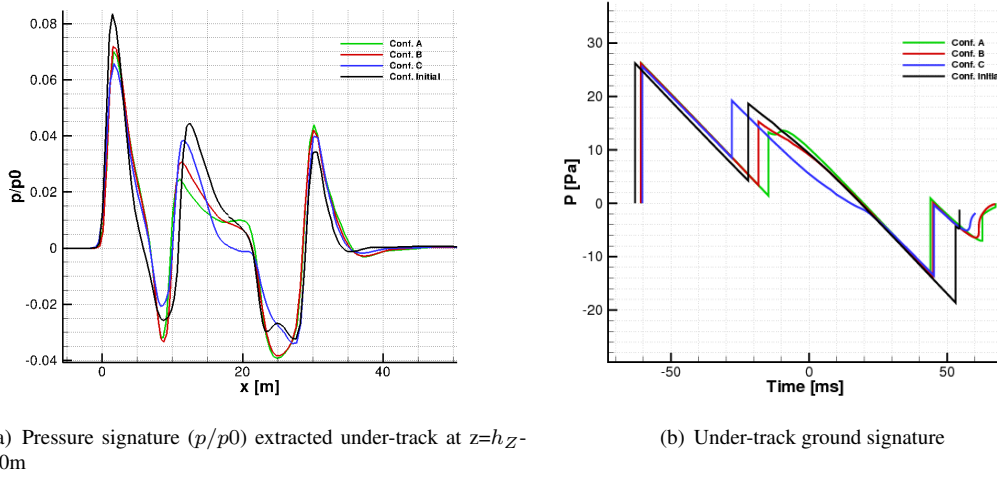


Figure 7: Near-field and ground pressure signal for different configurations on the Pareto front.

variables are able to act and shape significantly all the shocks and expansion waves. In particular all the configurations show a split of the rear shock due to the combined modification of the wing plan-form and rear fuselage. The initial peak in the near field p/p_0 is reduced of nearly 20 % with respect to the initial configuration, but this does not impact the ground level front shock pressure rise. In contrast the parameterization of the wing and of the rear part of the fuselage is able to produce modifications in the middle part of the near field signature. This corresponds to strong modifications on the ground signature pattern. As the sonic boom performance is improved, the pressure expansion just before the first maximum peak in the near-field signature is reduced, while the wing shock is increased in amplitude. The consequence at ground is a reduced amplitude of the second shock amplitude (from A to C). More design variables in the nose region should increase for a better description of the shape, and the ability of the algorithm to reduce the front shock amplitude.

4 CONCLUSION : COUPLING COOPERATION AND COMPETITION

A natural extension consists in combining cooperation and competition phases in the exploration of the Pareto front [8]. The competitive algorithm has been presented in [9]. It consists of a continuous succession of Nash games formulated with a special set of transformed variables defined from the diagonalization of the local reduced Hessian of one of the two disciplines, considered as the primary discipline. Here, we adapt this procedure for purpose of coupling with the MGDA.

Let the two objective functions to be denoted J_1 and J_2 .

The process begins with the optimization of J_1 alone, assumed to be conducted to full convergence. The corresponding point is one extreme of the Pareto-front segment. The competitive algorithm permits to generate a set of designs associated in the function space to a path tangent to the Pareto front, since it preserves the optimality of J_1 to a second-order term in the continuation parameter ϵ .

At some point along this initial path, one may decide to initiate the MGDA in order to generate a new segment bringing back the current point to the Pareto set. At this new point, the Pareto-stationarity condition writes:

$$\alpha \nabla J_1 + (1 - \alpha) \nabla J_2 = 0 \quad (7)$$

for some easily-identified α . Then, one can initiate again the continuation procedure based Nash games with Hessian-based territory splitting, now applied to the following pair of objective functions:

$$J_A = \alpha J_1 + (1 - \alpha) J_2 \quad J_B = J_2 \quad (8)$$

As a result, a new path in the (J_1, J_2) space is generated, tangent to the Pareto front, since it preserves the stationarity of J_A to a second-order term in the continuation parameter ϵ , and so on.

After drag-minimization based on high-fidelity CFD, the above procedure has been applied to a meta-model, and the resulting design-points have been re-evaluated by CFD *a posteriori*. The following steps are performed (see Figure 9):

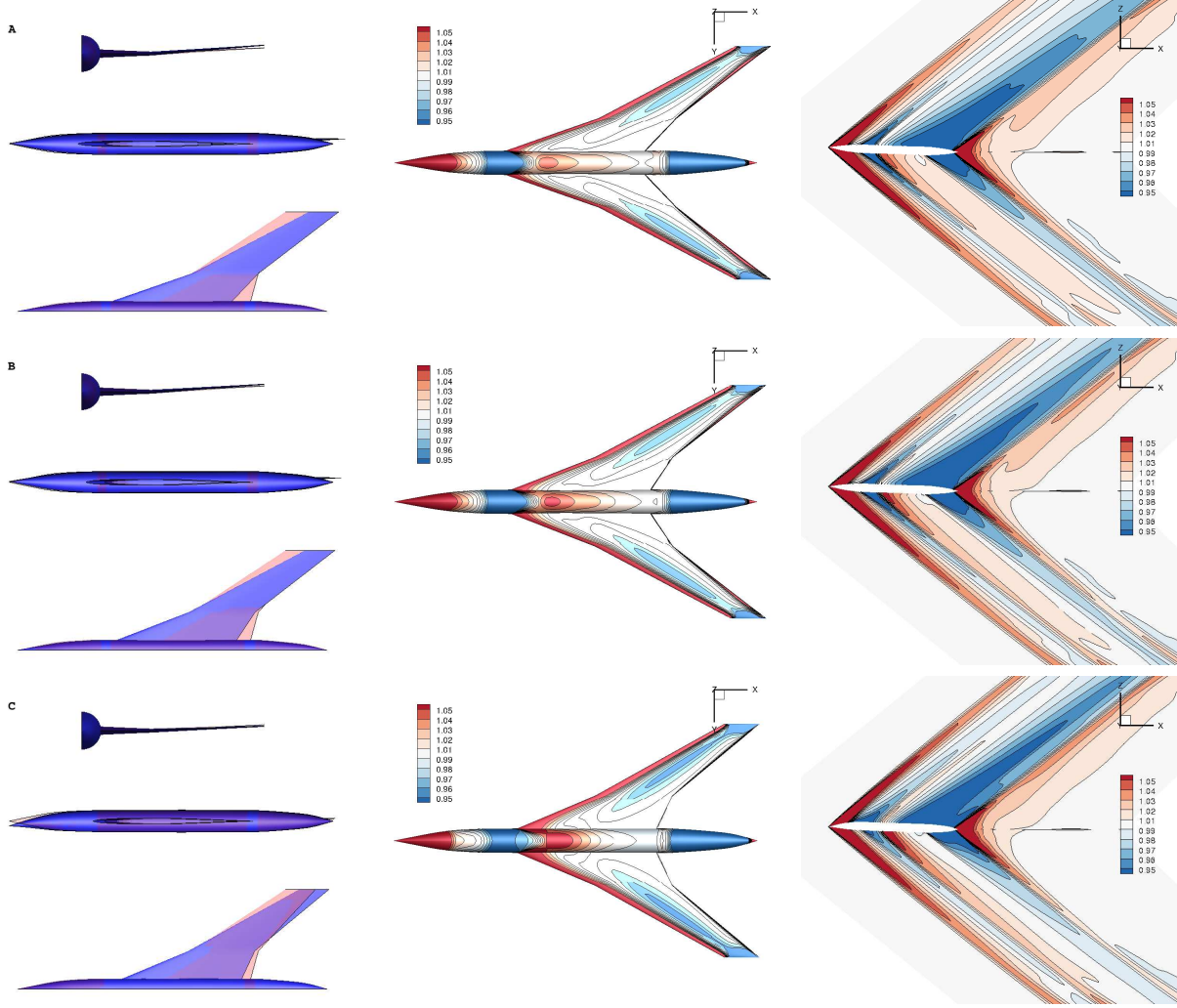


Figure 8: Geometry configuration, bottom skin pressure (p/p_0) and symmetry plane pressure (p/p_0) in near field for configuration A, B and C on the Pareto front (In red the initial geometry).

1. The minimization of the drag coefficient with constraint on lift and the competitive phase as in section [9] are performed providing respectively point 1 and 2;
2. A cooperative MGDA optimization is performed providing point 3 on the Pareto front;
3. A new α is evaluated in order to satisfy the Pareto-stationarity condition at 3;
4. A new split matrix is evaluated using the Hessian of the function $J_A(\alpha)$ where α has been evaluated at the previous point of the process;
5. Nash game is performed starting from the Pareto front in 3.

Points from 2 to 5 are repeated several times to provide additional solutions on the Pareto front.

In Figure 9 the Pareto front obtained using NSGA-II on a population of 24 individuals for 500 generations is also shown. Determining this front required approximatively ten times more function evaluations compared to the present cooperative-competitive method. In our experiment, each MGDA phase was initiated at a point naturally close to the Pareto front, and specifically after the criterion J_A (whose definition changes from a segment to the next), had been degraded by the succession of Nash games of at most 2%.

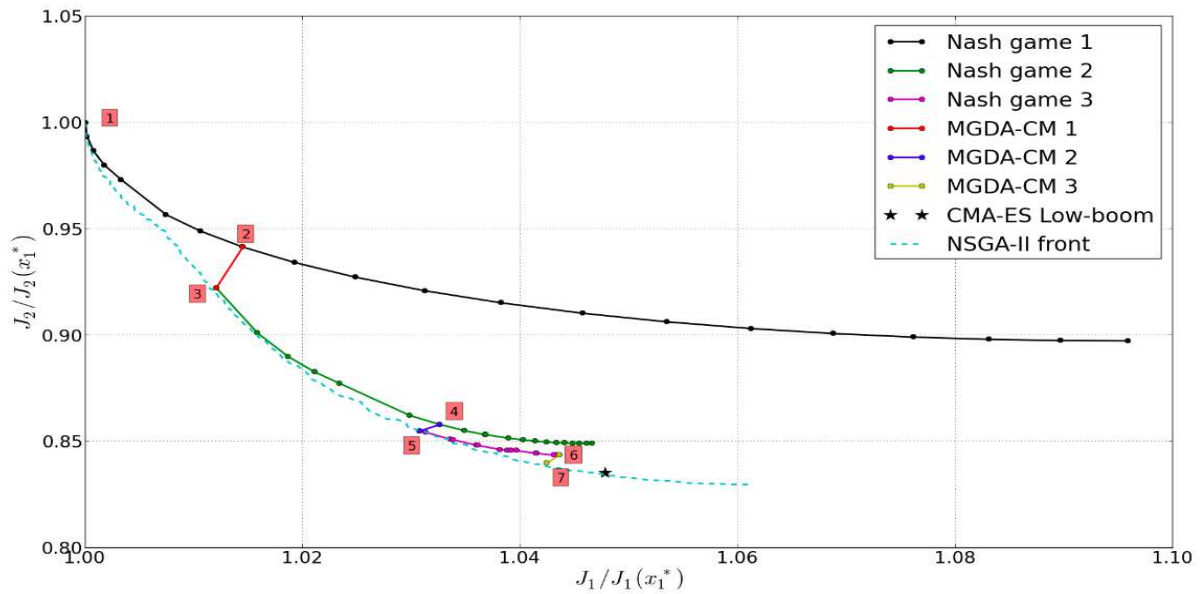


Figure 9: Coupling Nash games and MGDA iterations for the Pareto front exploration of the aero-acoustics problem.

References

- [1] J.A. Désidéri, *Multiple-Gradient Descent Algorithm (MGDA)*, Research Report 6953, INRIA, 2009, Revised version, Nov. 2012, <http://hal.inria.fr/inria-00389811>.
- [2] J.A. Désidéri, Application of MGDA to domain partitioning, Research Report 7968, INRIA, May 2012, <http://hal.inria.fr/hal-00694039>.
- [3] J.A. Désidéri, *MGDA II: A direct method for calculating a descent direction common to several criteria*, Research Report 7422, INRIA, April 2012
- [4] J.A. Désidéri, Multiple-gradient descent algorithm (MGDA) for multiobjective optimization, *Comptes rendus - Mathématique* 350 (2012), no. 5-6, 313-318, <http://dx.doi.org/10.1016/j.crma.2012.03.014>.
- [5] A. Zerbinati, J.A. Désidéri and R. Duveigneau: *Comparison between MGDA and PAES for multi objective optimization*, INRIA research report, number 7667. Sophia Antipolis June 2011, <http://hal.inria.fr/docs/00/60/54/23/PDF/RR-7667.pdf>
- [6] J.A. Désidéri, *Application of the Multiple-Gradient Descent Algorithm (MGDA) and Metamodels to a Multiobjective Problem in Aerodynamics*, *Proc. European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS)*, September 10-14, 2012, Vienna, Austria, 2012
- [7] K. Deb, L. Thiele, M. Laumanns and E. Zitzler, Scalable Test Problems for Evolutionary Multi-Objective Optimization. *TIK-Technical Report number 112*, Institut für Technische Informatik und Kommunikationsnetze, ETH Zürich Gloriastrasse 35., ETH-Zentrum, CH-8092, Zürich, Switzerland, 2001.
- [8] Minelli, A., Salah el Din, I., carrier, G., Zerbinati, A., and Désidéri, J.-A., "Cooperation and Competition Strategies in Multi-objective Shape Optimization Application to Low-boom/Low-drag Supersonic Business Jet", *AIAA Applied Aerodynamics Conference 2013*, 24-27TH June, San Diego CA.
- [9] J.-A. Désidéri, A. Minelli, and E. Roca Leon. "A cooperative algorithm for multi-objective optimization: Multiple Gradient Descent Algorithm (MGDA)", *4th Inverse Problems, Design and Optimization Symposium (IPDO-2013)* Albi, France, June 26-28, 2013.
- [10] P. E. Gill, W. Murray, and M. H. Wright, "Practical Optimization", Academic Press New York London, 1986.